The Review *of* Economics *and* Statistics

VOL. LXXIII | FEBRUARY 1991 | NUMBER 1

IS CONSUMPTION TOO SMOOTH? LONG MEMORY AND THE DEATON PARADOX

Francis X. Diebold and Glenn D. Rudebusch*

Abstract—Under common ARIMA representations of income, the permanent-income hypothesis predicts that the volatility of consumption should be *larger* than the volatility of unanticipated shocks to income; this prediction is not supported by the data. We examine whether this apparent excess smoothness of consumption is the result of the ARIMA representation's implicit restrictions on low-frequency dynamics. By using a generalized long-memory stochastic representation, we construct confidence intervals for the long-run impulse response of income in the absence of such low-frequency restrictions. These intervals are quite wide and include regions in which excess smoothness vanishes.

I. Introduction

IN the last decade, a large amount of macroeconomic research has been devoted to various aspects of the permanent income hypothesis (PIH) under rational expectations. While Euler equations from the simplest models (e.g., Hall (1978)) imply that consumption should follow an approximate random walk, it is generally agreed that the data instead indicate that variables other than lagged consumption appear to play a significant role in determining current consumption (i.e., consumption displays "excess sensitivity"). Moreover, recent work has stressed that, given the empirical result that income appears to be highly persistent, the PIH implies that changes in consumption should be larger than the innovations to income. This implication does not appear to accord with the data because movements in consumption are smaller than income innovations. This apparent excess smoothness of consumption, relative to PIH predictions with persistent income, has been labeled the "Deaton paradox."[1]

Numerous economic arguments have been advanced to explain the phenomena of excess sensitivity and excess smoothness. A partial listing of these explanations includes: liquidity constraints (Hall and Mishkin (1982)), nonconstant real interest rates (Michener (1984) and Hall (1988)), precautionary saving (Caballero (1990)), aggregation over time (Christiano, Eichenbaum, and Marshall (1987)), aggregation over agents (Deaton (1987)), transitory consumption (Flavin (1981, 1988)), divergence between the information sets of econometricians and economic agents (West (1988), Flavin (1988), and Quah (1990)), habit formation (Deaton (1987)), and non-separable utility functions (Campbell and Mankiw (1990)).

Many of the above research strategies are theoretical attempts to make the stylized PIH model more "realistic" by introducing various "real world" complications; however, none of these modifications has garnered wide support. Even more importantly, before introducing *economic* complications to the model in response to its alleged empirical failure, one should be sure that such empirical failure is not due to arbitrary *statistical* assumptions imposed when testing the model. In this vein, we investigate econometric issues related to the Deaton paradox.[2] Specifically, we examine the consequences of relaxing restrictions on the representation of the stochas-

Received for publication June 6, 1989. Revision accepted for publication March 30, 1990.

* University of Pennsylvania and Board of Governors of the Federal Reserve System, respectively.

We thank two referees and Fallow Sowell for useful comments, as well as seminar participants at the 1988 North American Winter Meetings of the Econometric Society, especially Ken West, and at the Board of Governors of the Federal Reserve System. The views expressed do not necessarily reflect those of the Board of Governors or its staff. Diebold gratefully acknowledges financial support from the National Science Foundation (Grant SES 89-21715) and the University of Pennsylvania Research Foundation (Grant 3-71441).

[1] Excess smoothness, first noted by Deaton (1987), is examined in Campbell and Deaton (1989), Campbell and Mankiw (1990), and West (1988), while excess sensitivity is discussed in Flavin (1981) and Campbell (1987). Some of the subtle connections between the two are considered in Campbell and Deaton (1989) and Flavin (1988).

[2] However, we do not attempt to explain the excess sensitivity of consumption or other inconsistencies between PIH predictions and the data.

tic process generating income. We use a long-memory, fractionally-integrated model that permits a wide range of low-frequency behavior and nests the ARIMA specification as a special case.

In section II we set forth the standard PIH model of consumption and describe the crucial role in the excess-smoothness paradox that is played by the long-run properties of the stochastic process generating income. The class of fractionally-integrated models is introduced and issues related to time-series representations of income are discussed in section III. Empirical results are contained in section IV, where the model with fractional integration highlights the uncertainty associated with estimates of the parameter linking income innovations to changes in consumption. Section V concludes.

II. The Permanent-Income Hypothesis

Suppose that an infinitely-lived representative agent at time t must choose consumption in period t, C_t, in the face of a stream of stochastic future real labor income payments, Y_{t+i}, $i = 0, 1, \dots, \infty$. The consumer assumes the real interest rate, r, will be constant over the infinite planning horizon and possesses an endowment of non-human wealth of W_t at the end of period t. We take the permanent income hypothesis to imply that the consumer will set consumption in period t equal to contemporaneous ex ante permanent income, y_t^p, which is the annuity value of non-human wealth and expected human wealth; thus,

$$C_t = y_t^p \equiv [r/(1+r)]\left[W_t + \sum_{i=0}^{\infty} \beta^i E_t Y_{t+i}\right] \tag{1}$$

where $\beta \equiv 1/(1+r)$ and E_t is the operator for expectations formed at time t.[3] The evolution of wealth over time is given by

$$W_t = (1+r)(W_{t-1} + Y_{t-1} - C_{t-1}). \tag{2}$$

Following Flavin (1981), the first difference of equation (1) can then be written

$$\Delta C_t = r \sum_{i=0}^{\infty} \beta^i [E_t Y_{t+i} - E_{t-1} Y_{t+i}], \tag{3}$$

so that changes in consumption are driven by revisions in conditional expectations of future labor income.

Under rational expectations, the nature of the process generating labor income determines the behavior of consumption.[4] The evaluation of PIH predictions of consumption behavior then hinges on the appropriate specification of the income process. Deaton (1987) formulated the excess smoothness paradox in conjunction with accumulating empirical evidence (e.g., Nelson and Plosser (1982)) that many macroeconomic variables are well characterized as having unit roots. Specifically, for the case of real labor income, Campbell (1987) is able to find no evidence against the unit root null hypothesis. Assuming a unit root, the generating process for real income can be written as

$$\Phi(L)\Delta Y_t = \gamma' + \Theta(L)\epsilon_t, \tag{4}$$

where

$$\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$$
$$\Theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_q L^q,$$

and all roots of $\Phi(z)$ and $\Theta(z)$ are outside the unit circle.[5] Flavin (1981) and Hansen and Sargent (1981) provide the technology needed to compute the sequence of revisions in expected incomes,

$$\{E_t Y_{t+i} - E_{t-1} Y_{t+i}\}_{i=0}^{\infty},$$

and hence the change in consumption (3), following a shock ϵ_t to income. In fact, the change in consumption is proportional to the income shock,

$$\Delta C_t = \kappa \epsilon_t, \tag{5}$$

[3] Strictly, equation (1) is only true if the third and higher derivatives of the underlying utility function are equal to zero; otherwise, the uncertainty associated with future income flows may generate precautionary saving, so that consumption is less than permanent income. Caballero (1990) argues that such behavior could explain the excess smoothness phenomenon.

[4] We model the information set used by the consumer in projecting income to include only lags of income. As noted by West (1988), the effect of additional information could make consumption appear to be too smooth relative to the innovations of a univariate income process; however, his results indicate that this is unlikely to be the case. Campbell and Deaton (1989) reach a similar conclusion on this issue.

[5] It is not our intent here to address the "trends vs. unit roots" debate in the context of labor income, as in Christiano (1987), or, more generally, in Rudebusch (1990). Rather, as will be discussed subsequently, we work in a "stochastic trend" environment, but we explicitly broaden the analysis to include forms of long memory other than unit roots.

where

$$\kappa \equiv \frac{1 - \sum_{i=1}^{q} \beta^i \theta_i}{1 - \sum_{i=1}^{p} \beta^i \phi_i} \tag{6}$$

This key relationship provides the link between the stochastic properties of income (4), as captured in κ and ϵ_t, and the stochastic properties of consumption.

It will prove useful to rewrite income in the (equivalent) moving-average form,

$$\Delta Y_t = \gamma + A(L)\epsilon_t, \tag{7}$$

where

$$A(L) = \Phi^{-1}(L)\Theta(L) \equiv 1 + a_1 L + a_2 L^2 \cdots$$
$$\gamma = \Phi^{-1}(1)\gamma'.$$

Then we have as well

$$\Delta C_t = \kappa \epsilon_t, \tag{8}$$

where

$$\kappa = \left(1 + \sum_{i=1}^{\infty} \beta^i a_i\right) \equiv c_\infty^\beta. \tag{9}$$

The multiplier c_∞^β, which relates income innovations to changes in consumption, is simply the infinite cumulative impulse response, $A(1)$, adjusted to reflect the discount factor β. In other words, c_∞^β is the discounted sum of income changes resulting from the shock ϵ_t. If $r = 0$, $c_\infty^\beta = c_\infty^1 \equiv A(1)$, but if $r > 0$, then c_∞^β can be greater or smaller than c_∞^1, depending on the entire shape of the cumulative impulse response function.

Taking the standard deviation of each side of the consumption response equation (8) yields

$$\text{std}(\Delta C_t) = c_\infty^\beta \, \text{std}(\epsilon_t). \tag{10}$$

Deaton (1987) and Campbell and Deaton (1989) show that a variety of ARIMA specifications for the real income process (together with reasonable assumptions regarding the real interest rate) lead to the same conclusion: c_∞^β is substantially above unity.[6] Thus, under the PIH, the variability of observed consumption changes should be greater than the variability of income innovations; in fact, the opposite is observed, as consumption appears to be too smooth. For example, Deaton (1987) finds that the standard deviation of the growth of consumption is only about half of the innovation standard deviation of an ARIMA (1, 1, 0) process describing income.

In the PIH consumption model sketched so far, the underlying result is that changes in consumption depend upon the process generating income. The modeling of the income process is thus crucial for interpreting consumption behavior and evaluating the PIH. In the next section, we introduce a more general approximation to the Wold representation than those used previously, in order to more closely examine the low-frequency properties of income.

III. Allowing for Fractionally-Integrated Income

As in Diebold and Rudebusch (1989), we consider a generalization of the standard ARIMA (p, d, q) model to allow fractional integration:

$$\Phi(L)(1 - L)^d Y_t = \Theta(L)\epsilon_t, \; \epsilon_t \sim (0, \sigma_\epsilon^2), \tag{11}$$

where

$$\Phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p,$$

$$\Theta(L) = 1 - \theta_1 L - \cdots - \theta_q L^q,$$

all roots of $\Phi(L)$ and $\Theta(L)$ lie outside the unit circle, and d is allowed to assume values in the real, as opposed to the integer, set of numbers.[7] Stationarity and invertibility require $|d| < 1/2$, which can always be achieved by taking a suitable number of differences. Econometricians typically have considered only integer values of d; the leading special cases are the discrete options $d = 0$ and $d = 1$. However, allowance for non-integer d values, that is, fractional integration, provides for parsimonious yet flexible modeling of low-frequency variation. Operationally, a binomial

[6] It is not the unit root assumption per se that is responsible for this conclusion. Instead, the short memory dynamics that result for disposable income under the assumption of a unit root (that is, positive serial correlation in first differences) work to produce an impulse response that is greater than one.

[7] Fractional integration allows a local generalization of the unit root hypothesis; rather than forcing $d = 1$, we allow $1/2 < d < 3/2$.

pansion of the operator $(1-L)^d$ is used:

$$(1-L)^d = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)L^j}{\Gamma(-d)\Gamma(j+1)} \tag{12}$$

$$= 1 - dL + \frac{d(d-1)}{2!}L^2 - \frac{d(d-1)(d-2)}{3!}L^3 + \cdots, \tag{13}$$

where $\Gamma(\cdot)$ denotes the gamma, or generalized factorial, integral. When $d = 1$, this is just the usual first-differencing filter. For non-integer d, the operator $(1-L)^d$ provides an infinite-order lag-operator polynomial with coefficients that decline very slowly. We denote this general representation, with potentially fractional d, as the ARFIMA (AutoRegressive Fractionally-Integrated Moving Average) model.

The ARFIMA model can be put in the moving-average form used above to calculate the present discounted value of future income. First, write equation (11) as

$$(1-L)^d Y_t = B(L)\epsilon_t, \tag{14}$$

where $B(L) = \Phi^{-1}(L)\Theta(L)$. Extracting the factor $(1-L)$ gives

$$(1-L)^{d-1}(1-L)Y_t = B(L)\epsilon_t, \tag{15}$$

or

$$(1-L)Y_t = A(L)\epsilon_t, \tag{16}$$

where $A(L) = (1-L)^{1-d}B(L)$.[8]

The ARFIMA model (11) belongs to the class of long-memory processes, so-named for their ability to display significant dependence between observations widely separated in time.[9] Standard ARMA processes are short-memory because the autocorrelation (or dependence) between observations τ periods apart decays rapidly as τ increases. Indeed, the autocorrelations at lag τ decline exponentially:

$$\rho_Y(\tau) \sim r^\tau, \quad 0 < r < 1, \tau \to \infty.$$

For ARFIMA processes, however, the autocorrelation function has a much slower hyperbolic decline:

$$\rho_Y(\tau) \sim \tau^{2d-1}, \quad d < .5, \tau \to \infty.$$

The intuition of long memory and the limitation of the integer-d restriction emerge clearly in the frequency domain. The series $\{Y_t\}$ displays long memory if its spectral density, f_Y, increases without limit as angular frequency tends to zero:

$$\lim_{\lambda \to 0} f_Y(\lambda) = \infty. \tag{17}$$

In fact, for an ARFIMA series, $f_Y(\lambda)$ behaves like λ^{-2d} as $\lambda \to 0$, so d parameterizes the low-frequency behavior. This is in contrast to the usual ARIMA model, where the spectral density is forced to behave like λ^{-2} as $\lambda \to 0$. Thus, a rich range of spectral behavior near the origin becomes possible when the integer-d restriction is relaxed. The "typical spectral shape" of economic variables (Granger (1966)), is monotonically declining with frequency, except for possible peaks at seasonals, with high power at low-frequencies. This shape is well-captured by fractionally integrated processes. For example, while *levels* of economic data typically have high power at low frequencies, *differences* often have little power at low frequencies; this is characteristic of fractionally-integrated processes with $d < 1$.[10]

The ARFIMA model, by allowing a variety of spectral shapes near the origin (corresponding to the continuum of possible d values), can provide superior approximations to the Wold representations of economic time series. This is particularly important in the context of the Deaton paradox; as discussed in section II, an assessment of the excess smoothness of consumption depends critically on the estimates of the discounted sum of coefficients in a Wold representation.

IV. Empirical Results

A. Estimation of d

We use a two-step procedure for the estimation of fractionally-integrated models due to

[8] As in section II, we can allow for drift, $(1-L)Y_t = \gamma + A(L)\epsilon_t$, and do so in the estimation reported below.

[9] Granger and Joyeux (1980) and Hosking (1981) independently proposed the use of ARFIMA processes as flexible long-memory models. Further discussion of fractional integration can be found in Diebold and Nerlove (1990).

[10] Further intuition for the ARFIMA representation is provided by Granger (1980), who shows that fractional integration may be induced by aggregation. Specifically, if the underlying components of an aggregate series (e.g., firms' productions) follow AR(1) processes with parameters ρ_i, and the ρ_i's are beta-distributed in the cross section, then aggregation yields a fractionally-integrated macroeconomic series.

Geweke and Porter-Hudak (GPH) (1983). We first obtain a consistent and asymptotically normal estimate of d and transform the series by the expansion of $(1 - L)^d$. We then fit an ARMA model to the transformed series to obtain consistent estimates of the remaining model parameters Φ, Θ, and σ_ϵ^2. Finally, these estimates are used to construct consistent estimates of the discounted cumulative impulse response.

The first-stage estimate of d is based on the slope of the spectral density function near $\lambda = 0$. Denote the first difference of the relevant series, $X_t = (1 - L)\, Y_t$; we wish to estimate $\tilde{d}$ in the model

$$(1 - L)^{\tilde{d}} X_t = \Phi^{-1}(L)\Theta(L)\epsilon_t \equiv u_t. \tag{18}$$

As d of the level series equals $1 + \tilde{d}$, a value of $\tilde{d}$ equal to zero corresponds to a unit root in Y_t.

The spectral density of X_t is given by

$$\begin{aligned} f_X(\lambda) &= |1 - \exp(-i\lambda)|^{-2\tilde{d}} f_u(\lambda) \\ &= [2\sin(\lambda/2)]^{-2\tilde{d}} f_u(\lambda) \end{aligned} \tag{19}$$

where $f_u(\lambda)$ is the spectral density of the stationary process u_t. Suppose that a sample of size T is available $(X_t, t = 1, \ldots, T)$; let $\lambda_j = 2\pi j/T$ $(j = 0, \ldots, T - 1)$ denote the harmonic ordinates of the sample. Taking logarithms of equation (19), adding and subtracting $\ln\{f_u(0)\}$, and evaluating at the harmonic ordinates, we obtain

$$\begin{aligned} \ln\{f_X(\lambda_j)\} &= \ln\{f_u(0)\} - \tilde{d}\ln\{4\sin^2(\lambda_j/2)\} \\ &\quad + \ln\{f_u(\lambda_j)/f_u(0)\}. \end{aligned} \tag{20}$$

If we restrict consideration to the low-frequency ordinates near zero, say, $\lambda_j, j \leq K \ll T$, the last term in (20) can be dropped as negligible. Let $I(\lambda_j)$ denote the periodogram at ordinate j, then add $\ln\{I(\lambda_j)\}$ to both sides of (20) and rearrange to obtain

$$\begin{aligned} \ln\{I(\lambda_j)\} &= \ln\{f_u(0)\} - \tilde{d}\ln\{4\sin^2(\lambda_j/2)\} \\ &\quad + \ln\{I(\lambda_j)/f_X(\lambda_j)\}. \end{aligned} \tag{21}$$

The particular utility of this formulation is its formal similarity to a simple linear regression equation:

$$\ln\{I(\lambda_j)\} = \beta_0 + \beta_1 \ln\{4\sin^2(\lambda_j/2)\} + \eta_j, \qquad j = 1, \ldots, K \tag{22}$$

where β_0 is the constant $\ln\{f_u(0)\}$, and the η_j, equal to $\ln\{I(\lambda_j)/f_X(\lambda_j)\}$, are independently and identically distributed.

Now let the number of low-frequency ordinates used in the above spectral regression be a function of the sample size, i.e., $K = g(T)$. Then, under regularity conditions on $g(\cdot)$, essentially that $g(T)$ approach ∞ with T, but at a slower rate, the negative of the OLS estimate of the slope coefficient provides a consistent and asymptotically normal estimate of $\tilde{d}$. This is true regardless of the orders and parameterizations of the Φ and Θ polynomials underlying the stationary process u_t. Furthermore, the variance of the estimate of β_1 is given by the usual OLS estimator, and the theoretical asymptotic variance of the regression error η_j was shown by Geweke and Porter-Hudak (1983) to be equal to $\pi^2/6$, which can be imposed to increase efficiency.

The regularity conditions on $g(T)$ required for a consistent estimate of d do not uniquely determine $g(T)$ or K, the sample size for the GPH regression. However, since d parameterizes and is estimated from the long-run dynamics of the time series, economic considerations can suggest a reasonable definition of the long run and hence designate a reasonable GPH sample size. In particular, we exclude from the estimation of d the information contained in short-run business cycles, namely, cyclical movements with periods of five years or less.[11] This implies that periodogram ordinates at frequencies *lower* than the five-year frequency should be included in the GPH regression, which translates into a "cutoff frequency" for inclusion in the regression of $2\pi/5$ for annual data and $2\pi/20$ for quarterly data. In the time domain, this five-year cutoff criterion can be translated into a simple and intuitive rule: the number of periodogram ordinates included in the GPH regression (K) should equal the number of non-overlapping five-year intervals available in the data sample.[12] There is a trade-off involved in specifying the period of the shortest cycle included. A shorter cutoff period gives a larger GPH regression sample and hence smaller stan-

[11] According to the National Bureau of Economic Research's chronology of turning points, the average length of a U.S. business cycle is 52 months.

[12] Formally, let p denote the cutoff period, that is, the length of the shortest cycle to be included (expressed in units equivalent to the frequency of measurement of the series). Then the corresponding cutoff frequency is $2\pi/p$. Now, the K^{th}, or final, periodogram ordinate is at frequency $2\pi K/T$. Equating the two gives $2\pi K/T = 2\pi/p$, or $K = T/p$.

TABLE 1.—ESTIMATES OF d FROM GEWEKE-PORTER-HUDAK REGRESSION

Data Series and Source	$\hat{d}$
Quarterly disposable income 1953:Q4–1987:Q2 (NIPA)	1.10 (0.38)
Annual disposable income 1929–1987 (NIPA)	1.17 (0.26)
Quarterly labor income 1953:Q4–1984:Q4 (Blinder and Deaton (1985))	0.99 (0.43)

Notes: All variables are in constant dollars on a seasonally adjusted per capita basis. The standard errors given in parentheses are constructed imposing the known theoretical regression error variance of $\pi^2/6$. The sample size used in the GPH regression is $T/20$ for quarterly series and $T/5$ for annual series, rounded to the nearest integer, as discussed in the text.

dard errors for the estimate of d; however, the shorter cutoff period also biases the estimate of d with high frequency information. Our results, however, were robust across a range of cutoff periods.

Table 1 provides GPH estimates of d for two income series: real disposable income, as reported in the National Income and Product Accounts (NIPA), and real labor income, as calculated by Blinder and Deaton (1985, table A-1).[13] The first series, although commonly used in testing the permanent income hypothesis, is somewhat too broad, as it includes all sources of income and not just income from labor. The Blinder and Deaton series, on the other hand, is an approximate labor income series equal to wages and salaries and a weighted sum of several other categories from the NIPA breakdown of personal income. In short, the two series have contrasting costs and benefits: the disposable income series is too broad but is directly measured, while the labor income series is conceptually more appropriate but judgmentally constructed. We therefore examine both; we estimate the fractional integration parameter for the per capita level of quarterly disposable income and labor income.[14] We also examine the NIPA annual disposable income series, for which a much longer time span of data (from 1929 through 1987) is available.

The point estimates of d in table 1 are all in the range of 1.0; however, it is important to note their large associated standard errors. Two standard deviations from the point estimates encompasses the range from 0.5 to 1.5; thus, the confidence we can have in the estimates of d is quite low. Because of these results, we are wary of *conditioning* upon the assumption of a unit root and proceeding to estimate models based on first-differenced data.[15] Instead, we perform a full sensitivity analysis, explicitly acknowledging the uncertainty associated with any estimate of d. It has been suggested by others (e.g., Deaton (1987)) that the limited information in the available macroeconomic data precludes a determination of the long-run properties of labor income. We are able to provide explicit, formal support for this thesis below, as the discounted cumulative impulse response is seen to depend crucially upon the order of integration.

B. *Estimation of κ*

Computation of the discounted cumulative impulse response requires estimation of all of the parameters of the ARFIMA (p, d, q) model. In this section, we focus on quarterly real disposable income to compute this response and investigate its robustness to variations in the model.[16] We first transform the first difference of the level of disposable income by the long-memory filter (17) for a range of values of the fractional integration parameter. These transformed series are then modeled as ARMA (p, q) processes to capture the remaining short-run dynamics.[17] Finally, estimates of the discounted cumulative impulse response, conditional on a variety of assumed real interest rates, are constructed.

Tables 2 through 4 provide estimates of the discounted impulse response. The tables differ in their assumed discount factor β, with β equal to 1, 0.998, and 0.995 in tables 2, 3, and 4, respectively. These discount factors correspond to real

[13] Both data samples start in 1953:Q4 for conformity and to omit distortions during the Korean War.

[14] Campbell and Deaton (1989) have also formulated the Deaton paradox in the logarithm of income; we obtained similar d estimates with logs.

[15] This is particularly true in light of the low power of the usual unit root tests against fractionally-integrated alternatives (Diebold and Rudebusch (1990)).

[16] We focus on quarterly disposable income to conserve space, however, our results were similar for the other income series.

[17] Although the long-memory filter is truncated at each point to the available sample, the estimate of d from the GPH regression and the second-stage estimates of Φ and Θ are all consistent.

TABLE 2.—DISCOUNTED IMPULSE RESPONSE FOR QUARTERLY DISPOSABLE INCOME, $\beta = 1$

d	ARFIMA(p, d, q) (AIC) ARFIMA(p, d, q) (SIC)	c_{16}^{β}	c_{40}^{β}	c_{80}^{β}	c_{160}^{β}	c_{400}^{β}	c_{4000}^{β}
1.4	(2, 1.4, 0)	2.19	3.11	4.07	5.36	7.72	19.37
	(2, 1.4, 0)	2.19	3.11	4.07	5.36	7.72	19.37
1.2	(3, 1.2, 0)	1.70	2.03	2.33	2.68	3.22	5.10
	(0, 1.2, 0)	1.91	2.28	2.62	3.01	3.61	5.72
1.0	(3, 1, 3)	1.17	1.11	1.12	1.14	1.13	1.13
	(0, 1, 0)	1.00	1.00	1.00	1.00	1.00	1.00
0.8	(3, .8, 3)	1.17	0.91	0.80	0.71	0.58	0.37
	(3, .8, 0)	1.14	0.93	0.81	0.70	0.58	0.37
0.6	(3, .6, 3)	1.80	1.41	1.01	0.74	0.50	0.20
	(1, .6, 1)	1.70	1.43	1.03	0.74	0.50	0.20

Notes: For each d value, we report the models selected by the Akaike information criterion (AIC) and Schwartz information criterion (SIC). The ARFIMA (p, d, q) model, $(1 - L)^d Y_t = \text{ARMA}(p, q)$, is estimated conditional upon d. For each model, we report estimates of the discounted cumulative impulse response function, c_k^{β}, for $k = 16$, 40, 80, 160, 400 and 4000, as discussed in the text.

TABLE 3.—DISCOUNTED IMPULSE RESPONSE FOR QUARTERLY DISPOSABLE INCOME, $\beta = .998$

d	ARFIMA(p, d, q) (AIC) ARFIMA(p, d, q) (SIC)	c_{16}^{β}	c_{40}^{β}	c_{80}^{β}	c_{160}^{β}	c_{400}^{β}	c_{4000}^{β}
1.4	(2, 1.4, 0)	2.17	3.04	3.90	4.92	6.30	7.49
	(2, 1.4, 0)	2.17	3.04	3.90	4.92	6.30	7.49
1.2	(3, 1.2, 0)	1.69	2.01	2.27	2.55	2.87	3.09
	(0, 1.2, 0)	1.90	2.26	2.55	2.86	3.22	3.47
1.0	(3, 1, 3)	1.16	1.11	1.12	1.13	1.13	1.13
	(0, 1, 0)	1.00	1.00	1.00	1.00	1.00	1.00
0.8	(3, .8, 3)	1.17	0.93	0.83	0.75	0.68	0.65
	(3, .8, 0)	1.14	0.95	0.83	0.75	0.68	0.65
0.6	(3, .6, 3)	1.80	1.42	1.07	0.86	0.71	0.66
	(1, .6, 1)	1.69	1.44	1.09	0.85	0.71	0.65

Notes: See notes to table 2.

interest rates ranging from 0% to 2% per year. In each table, the order of fractional integration is varied across 0.6, 0.8, 1.0, 1.2, and 1.4. For each d value, we estimate ARMA models with up to three autoregressive parameters and three moving-average parameters. We distinguish these models through the Akaike and Schwarz information criteria (AIC and SIC, respectively), which are differentiated by their degrees-of-freedom adjustment of the maximized log-likelihood function. The tables report the two optimal models according to the information criteria for each d value.[18] For each of the ten ARFIMA (p, d, q) models in each table, the discounted cumulative impulse responses at a variety of economic horizons are presented.

Strictly speaking, the stylized model of section II assumed the consumer had a relevant economic horizon of infinity. In richer models, shorter horizons are also of interest.[19] For a variety of reasons, the consumer may be myopic and may take into account the persistence of income, say, only over the next four years (c_{16}^{β}), ten years (c_{40}^{β}), twenty years (c_{80}^{β}), forty years (c_{160}^{β}), one hundred years (c_{400}^{β}), or one thousand years (c_{4000}^{β}). All of these horizons are reported in tables 2 through 4.

Let us focus on the results in table 3, which are based on a representative discount rate of $\beta = 0.998$. Assuming a one-hundred-year horizon, the estimated multiplier from income innovations to changes in consumption (i.e., $\kappa = c_{400}^{.998}$), ranges from 0.71 for $d = 0.6$ to 6.30 for $d = 1.4$, which illustrates how wide the *interval* estimates are for the long-run response. To approximate $k\%$ con-

[18] The two models reported for each d value are illustrative; not surprisingly, given d, the long-run response is similar across the sixteen ARMA models fitted.

[19] The infinite horizon may represent infinitely-lived consumers or an operative bequest motive. In other models in which agents have finite but stochastic lifetimes or stochastic termination of generational linkages, the relevant horizon may be shorter.

TABLE 4.—DISCOUNTED IMPULSE RESPONSE FOR QUARTERLY DISPOSABLE INCOME, $\beta = .995$

d	ARFIMA(p, d, q) (AIC) ARFIMA(p, d, q) (SIC)	c_{16}^{β}	c_{40}^{β}	c_{80}^{β}	c_{160}^{β}	c_{400}^{β}	c_{4000}^{β}
1.4	(2, 1.4, 0)	2.15	2.94	3.67	4.38	5.03	5.20
	(2, 1.4, 0)	2.15	2.94	3.67	4.38	5.03	5.20
1.2	(3, 1.2, 0)	1.68	1.97	2.19	2.39	2.54	2.57
	(0, 1.2, 0)	1.89	2.21	2.46	2.68	2.85	2.89
1.0	(3, 1, 3)	1.16	1.12	1.12	1.13	1.13	1.13
	(0, 1, 0)	1.00	1.00	1.00	1.00	1.00	1.00
0.8	(3, .8, 3)	1.18	0.95	0.86	0.81	0.78	0.77
	(3, .8, 0)	1.15	0.96	0.87	0.81	0.77	0.77
0.6	(3, .6, 3)	1.78	1.44	1.14	0.99	0.92	0.91
	(1, .6, 1)	1.68	1.45	1.15	0.99	0.92	0.91

Notes: See notes to table 2.

fidence intervals for κ, we can vary d over its $k\%$ confidence range (obtained by exploiting the asymptotic normality of the first-stage GPH estimate) and condition upon estimated Φ and Θ values at each d value.[20] Clearly, varying the d estimate across two standard errors encompasses consumption volatility and smoothness relative to income innovations. As in Deaton (1987) and Campbell and Deaton (1989), the estimated income innovations from these fitted models appear to have a larger standard deviation than the standard deviation of changes in actual consumption. Unlike the previous work, however, we cannot reject such a response as a significant departure from the predictions of the PIH, that is, as *excess* smoothness of consumption, once the uncertainty associated with integration order is taken into account.

V. Conclusions

Under the permanent income hypothesis for consumption behavior, the long-run, or low-frequency, stochastic properties of income are decisive in determining the response of consumption to an innovation in income. Earlier ARIMA specifications have suggested that the response of consumption to news about income should be much larger than appears to be true in the data. In this paper, we explicitly estimate the (generally fractional) degree of integration, rather than conditioning on a particular d estimate (as is implicitly done with ARIMA representations). More importantly, we are able to explicitly acknowledge the uncertainty associated with d and trace the effects of that uncertainty through to the cumulative impulse response function of income. We find that, while our point estimates are similar to those of others, our interval estimates for the multiplier linking changes in consumption to income innovations are quite wide, reflecting the (previously unacknowledged) uncertainty associated with d. The wide confidence intervals underscore a fundamental econometric reality: precise inference about low-frequency behavior is very difficult given the short span of income data available.

Finally, it is important to note that our arguments do not provide a complete reconciliation of data and theory. Specifically, while we have shown that violations of the PIH due to excess smoothness may not be as important as was previously believed, we have not addressed failures of the PIH due to excess sensitivity. The fact that consumption responds to anticipated changes in income is unaffected by the methods that we have employed and remains inconsistent with the theory.

[20] This procedure provides a lower bound on the widths of the true confidence intervals, because it does not take into account the stochastic variation in the Φ and Θ estimates.

REFERENCES

Blinder, Alan S., and Angus Deaton, "The Time Series Consumption Function Revisited," *Brookings Papers on Economic Activity* (2, 1985), 465–511.

Cabellero, Ricardo, "Consumption Puzzles and Precautionary Savings," *Journal of Monetary Economics* 25 (1990), 113–136.

Campbell, John Y., "Does Saving Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis," *Econometrica* 55 (Nov. 1987), 1249-1274.

Campbell, John Y., and Angus Deaton, "Is Consumption too

Smooth?," *Review of Economic Studies* 56 (July 1989), 357–373.

Campbell, John Y. and N. Gregory Mankiw, "Permanent Income, Current Income, and Consumption," *Journal of Business and Economic Statistics* (July 1990), 265–279.

Christiano, Lawrence J. "Why Is Consumption Less Volatile than Income?," *Quarterly Review*, Federal Reserve Bank of Minneapolis (Fall 1987), 2–20.

Christiano, Lawrence, J., Martin Eichenbaum, and D. Marshall, "The Permanent Income Hypothesis Revisited," Research Dept. Working Paper No. 335, Federal Reserve Bank of Minneapolis (1987); forthcoming *Econometrica* (1991).

Deaton, Angus, "Life Cycle Models of Consumption: Is the Evidence Consistent with the Theory?," in T. F. Bewley (ed.), *Advances in Econometrics, Fifth World Congress*, Volume II (Cambridge: Cambridge University Press, 1987), 121–146.

Diebold, Francis X., and Glenn D. Rudebusch, "Long Memory and Persistence in Aggregate Output," *Journal of Monetary Economics* 24 (Sept. 1989) 189–209.

______, "On the Power of Dickey-Fuller Tests against Fractional Alternatives," FEDS Working Paper 119, Board of Governors of the Federal Reserve System (1990); forthcoming *Economics Letters* (1991).

Diebold, Francis X., and Marc Nerlove, "Unit Roots in Economic Time Series: A Selective Survey," in Thomas B. Fomby and George F. Rhodes (eds.), *Advances in Econometrics: Co-Integration, Spurious Regressions, and Unit Roots* (Greenwich, Connecticut: JAI Press, 1990), 3–69.

Flavin, Marjorie A., "The Adjustment of Consumption to Changing Expectations About Future Income," *Journal of Political Economy* 89 (1981), 974–1009.

______, "The Excess Smoothness of Consumption: Identification and Interpretation," manuscript, Department of Economics, University of Virginia (1988).

Geweke, John, and Susan Porter-Hudak, "The Estimation and Application of Long Memory Time Series Models," *Journal of Time Series Analysis* 4 (1983), 221–238.

Granger, Clive W. J., "The Typical Spectral Shape of an Economic Variable," *Econometrica* 34 (1966), 150–161.

______, "Long Memory Relationships and the Aggregation of Dynamic Models," *Journal of Econometrics* 14 (1980), 227–238.

Granger, Clive W. J., and R. Joyeux, "An Introduction to Long-Memory Time Series Models and Fractional Differencing," *Journal of Time Series Analysis* 1 (1980), 15–39.

Hall, Robert E., "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy* 86 (1978), 971–987.

______, "Intertemporal Substitution in Consumption," *Journal of Political Economy* (1988), 339–357.

Hall, Robert E., and Fredric S. Mishkin, "The Sensitivity of Consumption to Transitory Income: Evidence from Panel Data on Households," *Econometrica* 50 (1982), 461–481.

Hansen, Lars P., and Tom J. Sargent, "A Note on Weiner-Kolmogorov Prediction Formulas for Rational Expectation Models," *Economics Letters* 8 (1981), 266–260.

Hosking, J. R. M., "Fractional Differencing," *Biometrika* 68 (1981), 165–176.

Michener, R., "Permanent Income in General Equilibrium," *Journal of Monetary Economics* 13 (1984), 297–305.

Nelson, Charles R., and Charles I. Plosser, "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics* 10 (1982), 139–162.

Quah, Danny "Permanent and Transitory Movements in Labor Income: An Explanation for "Excess Smoothness" in Consumption," *Journal of Political Economy* 98 (June 1990), 449–475.

Rudebusch, Glenn D., "Trends and Random Walks in Macroeconomic Time Series: A Re-examination," Economic Activity Working Paper 105, Board of Governors of the Federal Reserve System (1990).

West, Kenneth D., "The Insensitivity of Consumption to News About Income," *Journal of Monetary Economics* 21 (1988), 17–33.